TRACKING A WILDFIRE IN AREAS OF HIGH RELIEF USING VOLUNTEERED GEOGRAPHIC INFORMATION: A VIEWSHED APPLICATION

by

Korey Robert Klein

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Geography

The University of Utah

May 2014

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Korey Robert Klein**

has been approved by the following supervisory committee members:

Thomas J. Cova	, Chair	**3/20/14** Date Approved
Philip E. Dennison	, Member	**4/8/14** Date Approved
Steven Farber	, Member	**4/6/14** Date Approved

and by **Andrea R. Brunelle** , Chair/Dean of

the Department/College/School of **Geography**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

During rapid-onset disasters, timely dissemination of warning information to the public is crucial. Official emergency information channels are often slow, leaving the public to monitor social media websites for more timely updates. Examining Twitter communications, or tweets, sent during the 2012 Waldo Canyon Fire, this research seeks to determine what level of descriptive information is sent through Twitter during a wildfire, whether or not that information can inform other users of changes in fire activity, and how the spatial and temporal information within a tweet can be used in conjunction with geographic information systems (GIS) to determine fire location and activity.

This research utilized geotagged tweets and viewshed analysis in GIS as a means of determining what portions of the wildfire are visible from each Twitter user. These visible areas, or viewsheds, were then overlapped with viewsheds from other users to generate shared viewsheds. Both individual and shared viewsheds were compared to the area of new fire growth to determine if burning areas could be more confidently identified by considering different user perspectives.

The shared viewshed method showed that while increasing the number of observations does result in a decrease in shared visible area, the portion of the shared viewshed that falls within the fire boundary significantly increases. Many groupings,

which were compiled based on time sent and ranged in size from two to eight tweets, could see more than 20% of the fire.

This research found that there is the potential for users to inform one another of changes in fire activity that may not be visible from different points of view. The addition of viewshed analysis adds another layer of valuable information to the tweets and could be useful if done in real-time, especially during events occurring at a smaller scale.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Background

Disasters occur in numerous forms and regularly threaten human populations at varying levels of vulnerability (Wisner, 2004). Often times, disasters occur in the form of rapid-onset events, in which preparation times are short and lives are immediately threatened (Smith, 2013). Recent examples include the single deadliest tornado in over a half century on May 21, 2011 in Joplin, Missouri (Simmons & Sutter, 2012) and the Black Saturday bushfires in the Australian state of Victoria in February 2009 (Cameron et al., 2009). During rapid-onset disasters such as these, the dissemination of warning information to the public, especially in a timely manner, is pivotal in saving lives (Sorensen, 2000).

Unfortunately, official information from emergency response organizations is often disseminated to the public at a slower rate than desired (Pultar et al., 2009). Recently, volunteered geographic information (VGI) through social media websites, such as Twitter or Facebook, has shown promise in filling the time-sensitive void of information during disasters (Goodchild & Glennon, 2010; Mileti et al., 2006; Palen et al., 2010; St. Denis et al., 2012; Starbird et al., 2012; Sutton et al., 2008; Verma et al., 2011). The rapidly growing number of location-aware devices has allowed citizens to function as sensors, reporting what they deem pertinent to the situation (Goodchild, 2007). Naturally, since the citizens providing VGI are not professionals operating under established constraints, the accuracy and validity of the information becomes a concern (Goodchild & Glennon, 2010). These concerns raise the question of whether geolocated VGI from multiple sources and locations can help track the movement of a rapidly advancing threat.

Accurately tracking the movement of a dynamic threat is not only crucial for emergency managers, but for the public as well. In the eyes of the public, when a threat is looming in their vicinity, the primary concern is likely whether their family or property is at risk. The answer generally is contingent on the threat's location, trajectory, and rate of progress. User-generated messages sent through Twitter may help disseminate information to answer the public's concerns, essentially allowing citizens to help one another address the lag in information distributed by authorities through official channels.

1.2 Study Aims

The goal of this research is to use geotagged Twitter communications—known as, "tweets"—from a wildfire event in conjunction with viewshed analysis to improve the understanding of the location and extent of an advancing fire. While viewshed analysis and VGI have been paired in research by Jones et al. (2013), these two topics have yet to be jointly applied to a natural disaster scenario. Tweets are limited to 140 characters, which can limit the depth of communicated information. When a user elects to geotag their tweet, which simply attaches geographic coordinates to their message, they are providing another layer of information without using any of the limited characters. The value of geotagged tweets increases when they are paired with viewshed analysis in a GIS, which is used to determine the area that is visible from a specific location on the

Earth’s surface (Ormsby & Alvi, 1999). In this research, viewsheds from individual tweets will be intersected to identify a shared visible area. The area will then be compared to the wildfire’s known boundary to determine if the shared viewshed can accurately depict the fire’s location. This research aims to answer the following questions:

1) What type of information can be extracted from tweets during a wildfire that may aid in understanding its location, movement, and attributes?
2) How can the spatial and temporal information within a tweet be used in conjunction with viewshed analysis to narrow the location, extent, and direction of an advancing wildfire in areas of high relief?
3) Does the spatial and temporal variation of tweets allow users at one location to inform users at another location of changes in wildfire location, extent, and behavior?

A description of Twitter and short review of the relevant literature is covered in Section 2, followed by a Section 3, which provides a description of the study area and event. Section 4 describes the data sources and provides a detailed methodology. The results and all associated figures and tables are presented in Section 5, with the discussion of the research questions and the closing remarks in Section 6 and 7, respectively.

2. LITERATURE REVIEW

2.1 Twitter

Twitter is a social media website that allows users to communicate with one another in 140-character or less messages, known as "tweets," which can be sent directly to another user or broadcast to a network of users. In addition to textual information, tweets can include photos, hashtags, and locational information. Hashtags, which utilize the '#' symbol, immediately followed by a relevant keyword or phrase, allow users to categorize their tweets. Once a tweet is sent, the hashtag becomes a clickable link that directs a user to a list of all other tweets using the same keyword or phrase. Users can also elect to include their locational information, or geotag, which imbeds the user's geographic coordinates or place-based location in the tweet.

2.2 VGI and Social Media in Disasters

Within the last 6 years, the emerging subfield of VGI, introduced by Goodchild (2007), has expanded rapidly. In his seminal article, Goodchild introduces the concept of humans as sensors, specifying that humans have the ability to compile relevant information around them and share it. In the first mention of disasters, he suggests that VGI allows for people to provide a better description of current conditions based on their familiarity with the area. This article directly led to further expansion of the VGI literature, especially concerning disasters. VGI applications to disasters have been described for wildfires (De Loungeville et al., 2009; Goodchild & Glennon, 2010; Pultar et al., 2009; St. Denis et al., 2012; Sutton et al., 2008), floods (Palen et al., 2010), and earthquakes (Sarcevic et al., 2012; Starbird & Palen, 2011).

Although the discussion of the inaccuracies associated with VGI are well described in the literature (Goodchild, 2007; Goodchild & Glennon, 2010; Mendoza et al., 2010; Sutten et al., 2008), it is also recognized for its potential to be utilized by official sources of information or even surpass the quality of those channels (Goodchild & Glennon, 2010; Pultar et al., 2009; St. Denis et al., 2012; Sutten et al., 2008). It is already becoming more common for emergency officials to play an active role in the dissemination of information on Twitter. Cooperation in the dissemination of fire spread and evacuation information has already been displayed by on-scene emergency officials and remote volunteers that posted the information on various social network websites (St. Denis et al., 2012). Remote volunteerism is one of the more creative applications of VGI in the realm of disasters and has been demonstrated in the 2010 Haitian Earthquake (Starbird & Palen, 2011; Zook et al., 2010) and the 2011 Shadow Lake fire (St. Denis et al.). Another unique application of VGI in disaster has been the use of natural language processing to extract situational awareness from tweets during disasters (Verma et al., 2011; Vieweg et al., 2010).

In perhaps the article most parallel to this research, De Loungeville et al. (2009) uses VGI to detect spatio-temporal data on forest fires in the South of France. While the authors used data from Twitter as well, of the 127 users that contributed to their final dataset, only 5 provided tweets with geographic coordinates (i.e., geotags). As a result, the location of most tweets had to be inferred from the context of the tweet or the generic location provided in the user's profile. By specifically looking at tweets tagged with geographic coordinates, this paper aims to increase the value of spatio-temporal data in tweets by adding another layer of detail using viewshed analysis.

2.3 Viewshed Analysis

In GIS, viewshed analysis uses stored elevation values, often from a digital elevation model (DEM), to determine the visibility based on the elevation of the observation point, as well as the surrounding area. In other words, a viewshed is determined solely by an area's topography, and does not account for obstructions such as trees or buildings. Despite the exclusion of physical obstructions, a viewshed analysis performed in an area of high relief still yields a more complex result than one performed in a relatively flat terrain.

Viewshed analysis was a prominent GIS research topic in the early 1990s. Some early applications of viewsheds used triangulated irregular network (TIN) models to represent the data (Goodchild & Lee, 1989), but that was quickly replaced by the digital elevation model (DEM) (Fisher, 1991; Fisher, 1992; Fisher, 1996). Much of Fisher's work focused on determining viewshed and DEM accuracy, something that Maloy and Dean (2001) also addressed by comparing computer-generated viewsheds with loosely determined viewsheds using in-situ photos. Viewshed analysis has occasionally been applied to natural hazards and emergencies. While Midkiff and Bostian (2002) applied viewshed analysis to determine the best location for deploying broadband internet towers during emergencies, the remainder of the applications in natural hazards focuses on wildfires, mainly concerning potential fire tower placement (Fisher, 1996; Goodchild & Lee, 1989; Pompa-Garcia et al., 2010).

2.4 GIS in Emergencies

Geographic information systems have served a key role in the four phases of emergency management: mitigation, preparedness, response, and recovery. Reviews of the applicability of GIS to these phases have been conducted by Cova (1999) and Cutter (2003). Applications in the mitigation phase are often associated with vulnerability assessments, which have been applied for many types of disasters, including floods (Messner & Meyer, 2006) and tsunamis (Wood & Good, 2004). GIS has also been used in determining social vulnerability to disasters (Chakraborty et al., 2005; Cutter & Emrich, 2006; Morrow, 1999).

GIS is primarily used to structure and implement emergency response plans, and as a result, the preparedness and response phases are often merged (Cova, 1999). GIS can also help in compiling information from multiple sources and scales into a single database capable of being utilized in mapping and decision-making. An area of emergency management that has emerged as the foremost application of GIS in these phases is evacuation planning. GIS models of evacuation plans have been developed for a range of hazards, with wildfire being the most relevant for this research (Cova et al., 2005; Cova & Church, 1997; Dennison et al., 2007; Pultar et al., 2009).

While the GIS applications of the first three phases of emergency management focus heavily on modeling, the initial period of the recovery phase often uses GIS to coordinate recovery activities, including the positioning of logistical support and resources and preliminary damage assessments (Cova, 1999; Cutter, 2003). GIS can also be used to communicate the progression of the response effort, such as the making of daily maps that illustrate the availability of services and resources for the public. Some of these GIS applications were used in the extensive recovery efforts of the United States' greatest disasters, including the terrorist events of September 11, 2001 and Hurricane Katrina in 2005 (Cutter, 2003; Mills, 2008).

3. STUDY AREA AND EVENT

The focus of this research is the Waldo Canyon Fire, which affected the Colorado Springs, Colorado area from June 23rd – July 10th, 2012. More specifically, the research focus is 1 day, June 26th, when dry conditions and high winds caused the fire to rapidly grow from 5,180 acres to just over 15,500 acres. On this day, the fire also encroached into the wildland-urban interface (WUI), forcing the evacuation of thousands of people and eventually destroying approximately 346 homes. At the time, the Waldo Canyon Fire became the most destructive and expensive fire in Colorado state history, with over 350 million U.S. dollars in insurance claims.

4. DATA AND METHODS

4.1 Outline

This research integrates user-generated data created via Twitter and GIS techniques, including viewshed analysis, in an effort to track the movement of an advancing wildfire. In the first subsection, the data sources are described in detail. The methodology is explained in the subsections that follow, with the extensive filtering of the Twitter dataset covered first. In order to perform a viewshed analysis, the DEMs of the study area had to first be manipulated, the process of which is described in the third subsection. This subsection also includes the generation and conversion of the viewsheds, and the description of the tools used to determine shared viewsheds and fire visibility follow in the last subsection.

4.2 Data Sources

The digital elevation models (DEMs) for this research were downloaded from the United States Geological Survey (USGS) National Map Viewer. The size and location of the study area required the download of four separate DEM files in an ArcGRID format, each of which covered a one square degree area with a resolution of 1/3 arc second (about 10 meters). To create the final DEM, the four rasters were mosaicked together and converted to a single raster with a 32-bit floating point pixel type, which covered an area ranging from 38 degrees to 40 degrees north and 104 degrees to 106 degrees west.

The wildfire boundary was generated using two shapefiles that were downloaded from the Geospatial Multi-Agency Coordination Group (GeoMAC) website. These shapefiles were drawn according to thermal infrared imagery at two different times, the first occurring before the rapid growth on June 26 (22:53 on June 25) and the second occurring the following night (01:30 on June 27). In order to focus on just the areas of new growth, all areas from the June 25 shapefile were removed from the June 27 shapefile, leaving the 10,000+ acre area that burned on June 26. The timing for this change in perimeter was confirmed using the MODIS Active Fire Detection maps.

The Twitter dataset was provided by colleagues at Floating Sheep, a collective of geography and Big Data researchers that originated at the University of Kentucky (Crampton et al., 2013). As a part of their DOLLY Project (Data On Local Life and You), Floating Sheep has been collecting every geotagged tweet worldwide since December 2011. The data that are captured with each geotagged tweet includes a tweet ID number; user ID number; the users profile bio; geographic coordinates for the tweet; a geotag and place type; the country, state, and county from which the tweet occurred; a timestamp; and the text and hyperlinks included in the tweet. Originally included in the dataset for this research were all geotagged tweets sent during the 18-day life of the fire (17, 481 tweets), but as the research focus was narrowed to the day with the greatest fire rate-of-spread (ROS), all tweets not sent on June 26 were removed, leaving a total of 1,302 tweets. These tweets encompassed a set area around the Waldo Canyon Fire perimeter, including Colorado Springs and the United States Air Force Academy. The filtering process for these data can be found in the next section.

4.3 Filtering of Twitter Data

Next, the tweets had to be further filtered to remove those with only place-based location. These are tweets whose geotag is given for the general coordinates of a place, which is usually a polygon and can range in size from a point-of-interest (POI) to a city. For this dataset, all of the place-based geotags were at the city level, which given its size and the inherent inaccuracy with using its generic coordinates, had to be removed. This step left 912 tweets, all of which included a geotag based on specific latitude and longitude coordinate pairs.

The 912 remaining tweets were run through a Python script that counted the number of occurrences for each word. Table 1 shows the word counts for the heuristically determined fire-related words, locations of interest, and photo hyperlinks that were then manually sorted to yield the final dataset. During the manual filtering, tweets that provided no description of the fire's activity or location were removed, as were tweets that included photos accompanied by phrases such as 'taken by a friend' or 'taken earlier', which indicate that the coordinates associated with the tweet do not match those of the photo. Furthermore, any tweets that were retweets, which act as a forward of a tweet posted by another user, were removed for the same reason. Of the tweets that remained, 82% (98) included photos. Each photo link was then opened in a web browser and deemed suitable for the research or not. Photos deemed unsuitable either pictured the smoke plume without any visual reference of the ground or contained views obstructed by homes. Lastly, the locations of the remaining photos were verified using Google Earth. In the end, 82 tweets were considered of value for this research, with 68 containing the event hashtag, 13 containing either 'fire', 'flames', or 'wildfire', and the last containing ‘Garden of the Gods’, which is a popular public park and tourist destination northwest of Colorado Springs.

4.4 DEM and Viewshed Manipulation

The four DEMs were imported into ESRI’s ArcMap, where they were converted to a new mosaic raster. Before generating viewsheds from each tweet location, two additional fields were applied to the twitter dataset: AZIMUTH1 and AZIMUTH2. Both azimuth fields are given directional values in degrees based on the location of each tweet. To explain the values for each field, consider the cardinal directions. Each observer is considered to have a 360-degree view. The value for AZIMUTH1 marks one cardinal direction, such as 180-degrees south, and the value for AZIMUTH2 marks a second cardinal direction, such as 0-degrees north. If an observer were to face in the direction of AZIMUTH1 and then rotate to their right until they were facing the direction of AZIMUTH2, everything visible between those two directions would be of concern. Thus, in this example, only the areas visible to the west of the observer would be included in the viewshed. Adding the azimuth fields allowed for the exclusion of portions of a viewshed in the opposite direction of the observer from the fire.

The viewsheds were then generated for each observer in each time group. In order for the viewsheds to include units of linear measurement, each required a series of GIS steps. Since the viewshed output is in raster format, the first step was to convert it to a series of polygons based on its value of visible or not visible. Once converted, all of the polygons with a value of 0, or not visible, were selected and removed, leaving only those polygons that represent visible areas. Lastly, the polygons were given a projected coordinate system in order for them to contain units of linear measurement. Based on the location of the study area and meters as the desired linear unit, NAD 1983 UTM Zone 13N was chosen.

4.5 Shared Viewsheds and Fire Visibility

After the viewsheds were projected, each one was separated into a small group based on the time the tweet occurred. The timeframe for the groupings were kept short, 20 minutes or less, due to the rapid spreading of the fire, which averaged just over 430 acres of new growth per hour that day. This method resulted in 30 temporal groupings, ranging in size from one to eight tweets, or observers, occurring in an average timeframe of just over 10 minutes. While there was a natural clustering of tweets throughout the day, 6 of the unique observers were included in more than one group. Due to the limited size of the final dataset, the decision to include these tweets in two groupings was made so that the shared viewshed methods could be tested against larger groups. Without this decision, the largest grouping would have consisted of only 6 observers. The resulting distribution of group sizes can be seen in the histogram in Figure 1.

The steps that follow the grouping of observers are diagrammed in Figure 2. Once the temporal groups were established, shared viewsheds were created for each combination of observers using the Intersect Tool in ArcMap. In order to determine which portions of the fire were visible, each shared viewshed was run through the Clip Tool, which removed any viewshed polygons that did not fall within the fire boundary. The total area statistics were then extracted from each shared viewshed, whether clipped or unclipped, and organized in a spreadsheet. In the spreadsheet, each total area was converted to acres, and the clipped and unclipped totals were mathematically compared to determine what percentage of the viewshed fell within the fire boundary, as well as the percentage of the fire covered. This directly ties back to determining the location and extent of the fire, with the theory that the more observers that see the area, the more likely it is that area is burning.

Table 1 – June 26 Tweet Word Counts

Word	Count
#WaldoCanyonFire	139
Fire	85
Colorado Springs	55
Smoke	30
Garden of the Gods	13
Wildfire	7
Flames	7

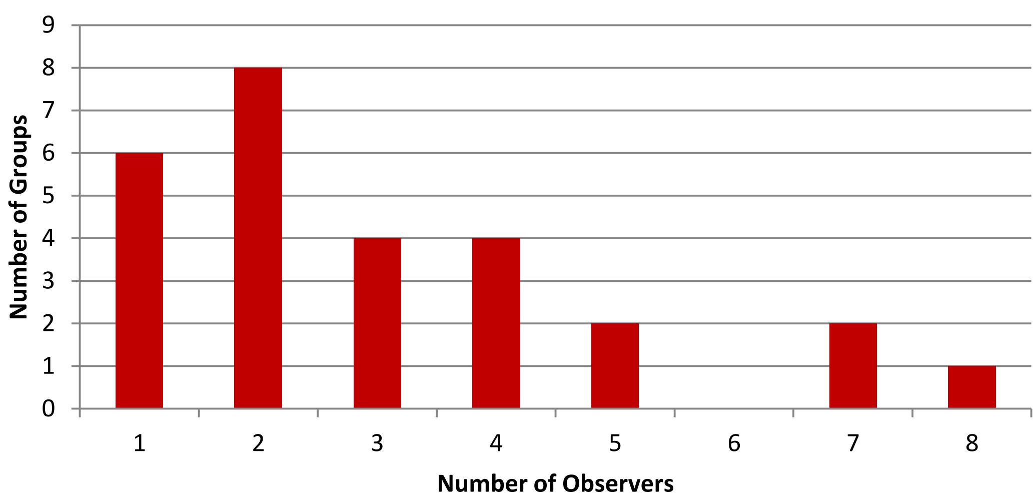

Figure 1 – Tweet Grouping Histogram

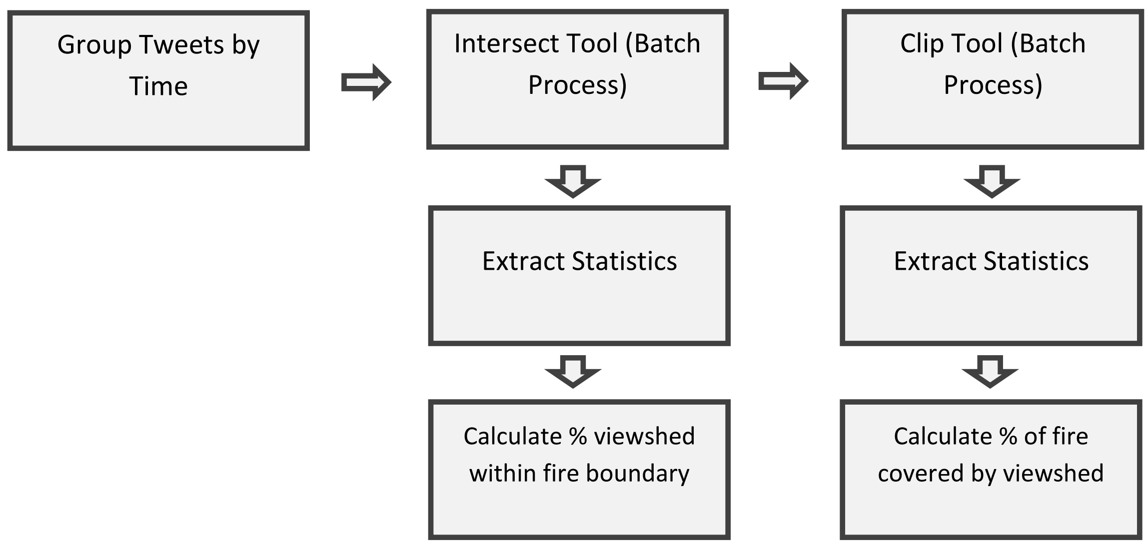

Figure 2 – Shared Viewshed and Result Calculation Process

5. RESULTS

For each grouping, viewsheds were generated for every observer and then intersected with each other to determine which areas were visible by a combination, or all, of the observers. Each of these shared viewsheds, as well as the individual viewsheds for each observer, was then clipped based on the fire boundary to determine which portions of the fire were visible. Figure 3 shows an example of the observer locations and corresponding viewsheds for a group of 4 observers from 6:19 pm to 6:33 pm. In this example, the 4 observers are distributed broadly, with the average distance between them being 13.6 kilometers. With such high spatial separation, it was expected that each observer's location would offer a unique vantage point, thus resulting in viewsheds that cover different portions of the fire.

Figure 4 illustrates the changes in the shared viewsheds for the same grouping as each observer is included. As can be seen by comparing Figure 4a and 4b, the addition of Observer 3's northern perspective to those of Observer 1 and Observer 2 drastically reduces the shared viewshed, with the total coverage dropping from about 4,376 acres to 882 acres. Conversely, the addition of the last perspective from Observer 4 in Figure 4c hardly alters the shared viewshed, only decreasing the total coverage by another 40 acres, bringing the total coverage visible by all 4 observers to 842 acres. This is likely due to Observer 4 being located only 4.55 kilometers from Observer 1, the closest distance between observers in this group, and thus not offering a greatly differing perspective.

Figures 5a, 5b, and 5c focus on the change in coverage strictly within the fire boundary, which follows a similar, slightly less drastic pattern, with coverage dropping from about 825 acres to 224 acres when adding Observer 3 and only decreasing an additional 3 acres when adding Observer 4. Even though this equates to a decrease from 8.0% to 2.2% to 2.1%, respectively, of the fire covered, the smaller factor of decrease in coverage inside the fire boundary versus the entire coverage means that a higher percentage of the shared viewshed falls within the fire boundary, an increase from 18.9% to 25.4% to 26.3%, respectively.

In addition to the above example, the same statistics were also calculated for each of the viewsheds in the other groups, and then consolidated across the entire day based on the number of observers, and organized into Table 2. For each unique combination of observers, two percentages are calculated: 1) percentage of viewshed within the fire boundary, and 2) percentage of fire covered by that viewshed. The median percentages are then determined for all unique combinations with the same number of observers across all groupings. Figure 6 shows how these percentages change when observers are added.

In Table 3, the total area of the fire visible from at least 1 observer was calculated for each group with more than 1 observer, as was the average distance between the observers. Of those groups, nine had 4 or greater observers and 12 had 2 or 3 observers. All of the groups with 4 or more observers could see more than 20% of the fire between them. While the largest group (8) did also have the largest percentage of the fire covered, the next three highest percentages came from groups of 4, not groups of 5 or 7.

Looking at the groups with 2 or 3 observers, 2 of the 12 groups could see less than 6% of the fire between them. On the other hand, five groups could see more than 17% of the fire between them, with the 5:13 pm to 5:28 pm group of 2 having the highest percentage, at 24%, as well one of the highest average distances between observers, at 16.6 kilometers. The two preceding groupings (4:20 – 4:25 pm and 4:38 – 4:45 pm) also have 2 observers, as well as large spatial distribution; however, they cover two vastly different percentages of the fire at 5% and 17%, respectively.

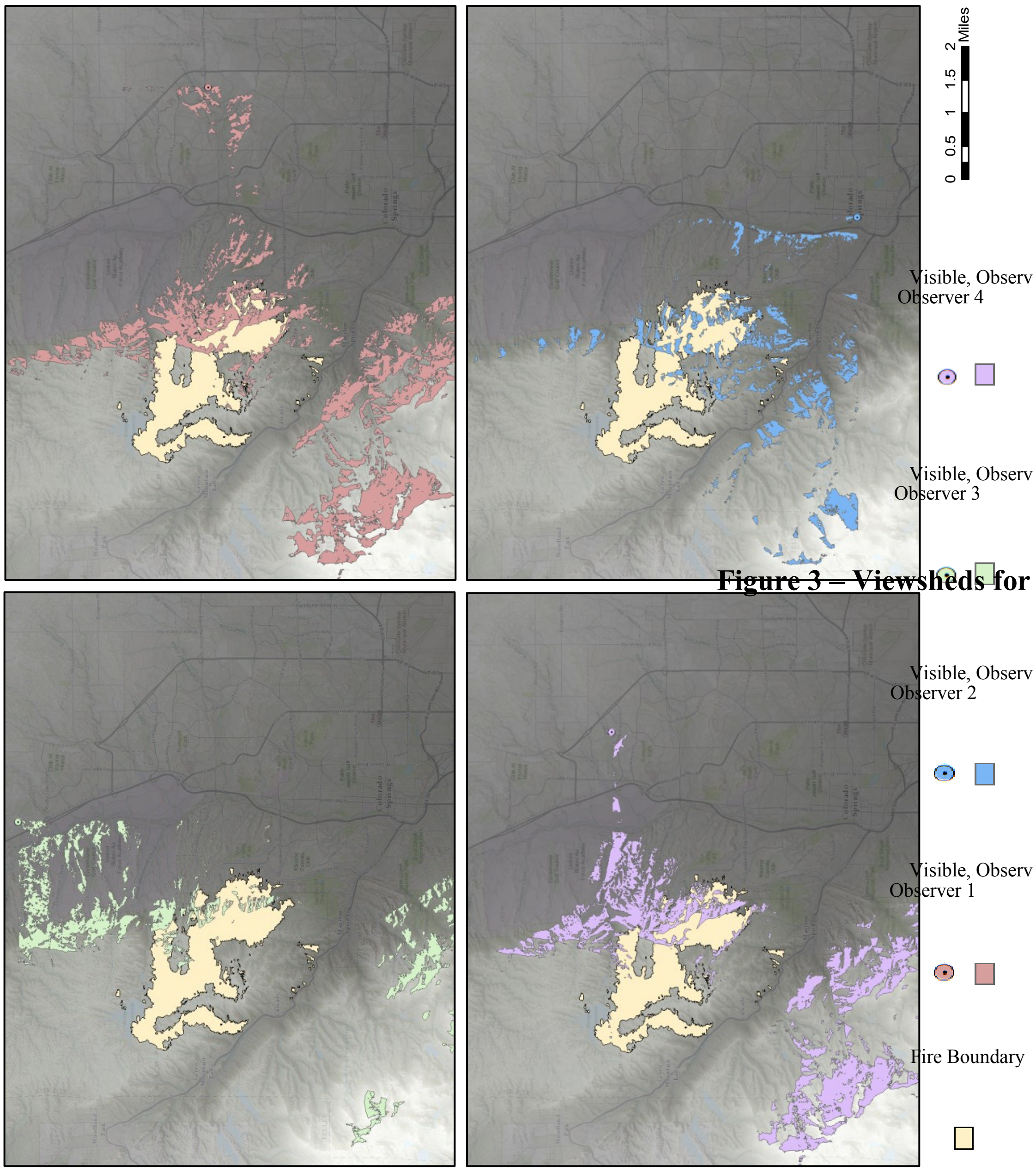

Figure 3 – Viewsheds for 4 Observers, 6:19 – 6

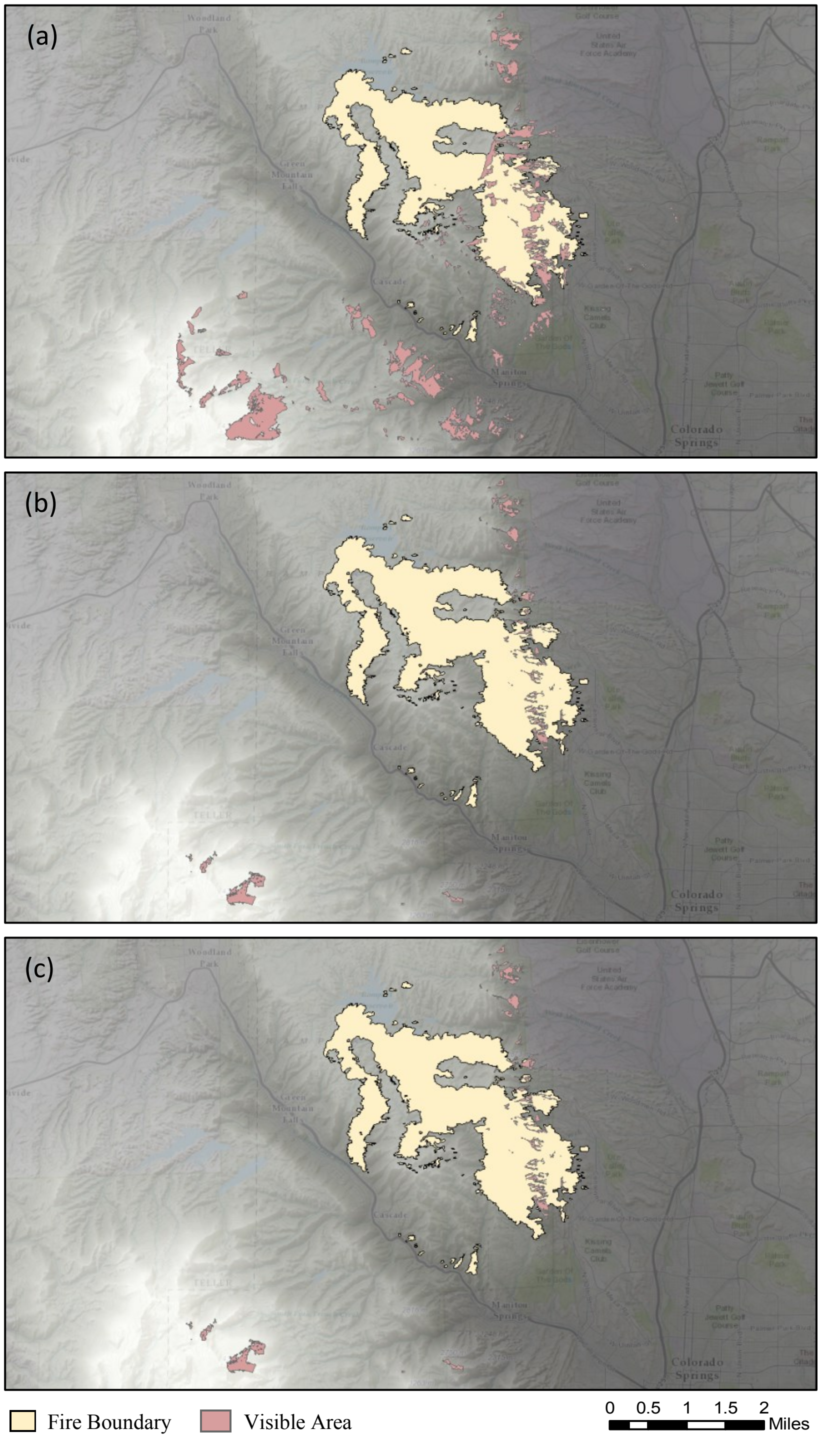

Figure 4 – Change in Shared Viewshed

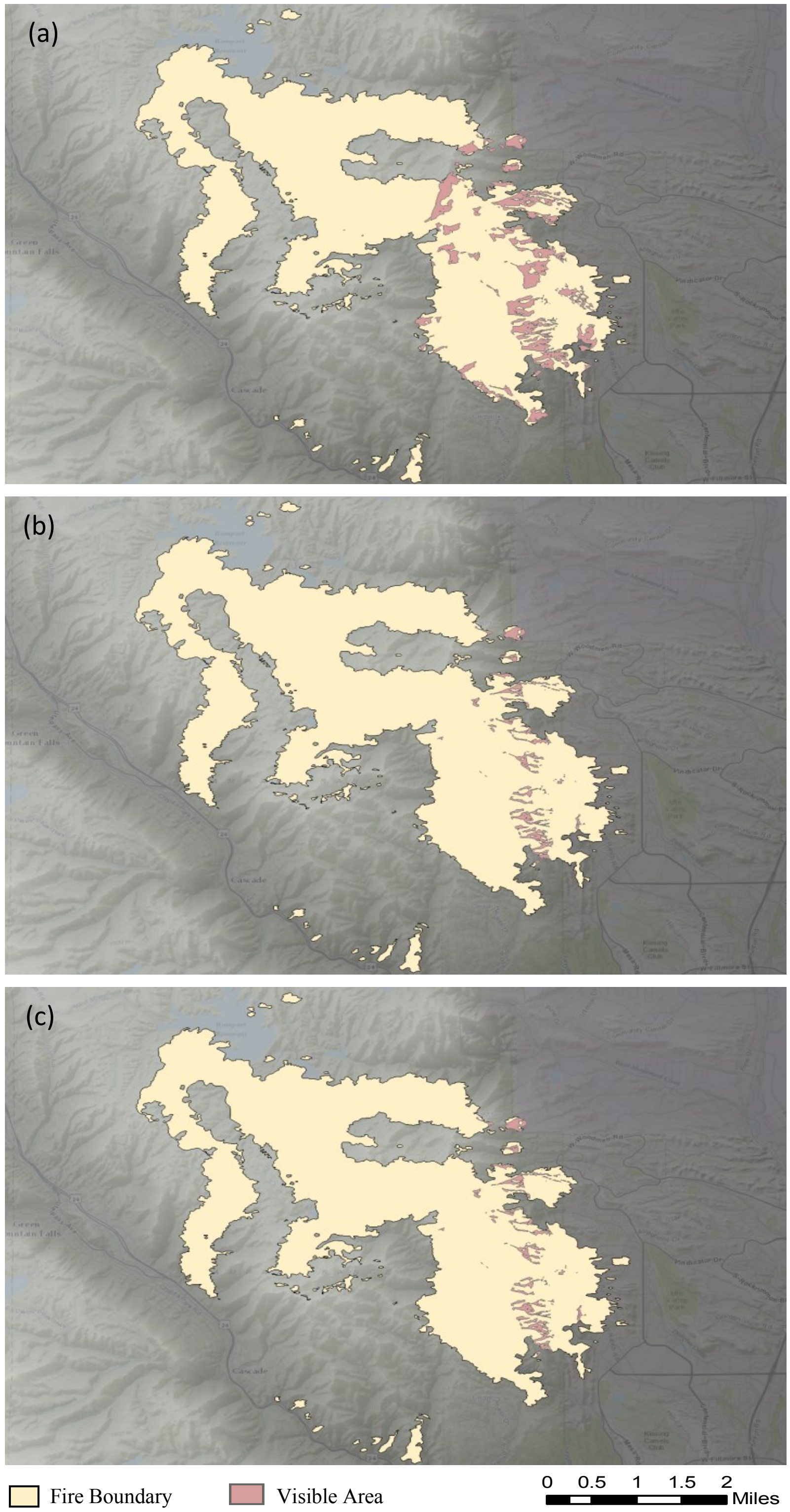

Figure 5 – Change in Shared Fire Visibility

Table 2 – Statistics for all Observer Groupings of Same Size

Number of Observers	Number of Unique Groupings	Average Coverage within Fire Boundary (acres)	Median Coverage within Fire Boundary (acres)	Median Percentage within Fire Boundary	Median Percentage of Burned Area Covered
1	85	1,177.5	1,460.4	4.9%	14.1%
2	134	652.4	497.1	6.5%	4.8%
3	166	394.2	223.4	7.2%	2.2%
4	154	225.9	148.5	19.3%	1.4%
5	100	130.4	74.2	38.6%	0.7%
6	42	82.0	660	75.4%	0.6%
7	10	62.0	57.8	76.6%	0.6%
8	1	56.9	56.9	76.6%	0.6%

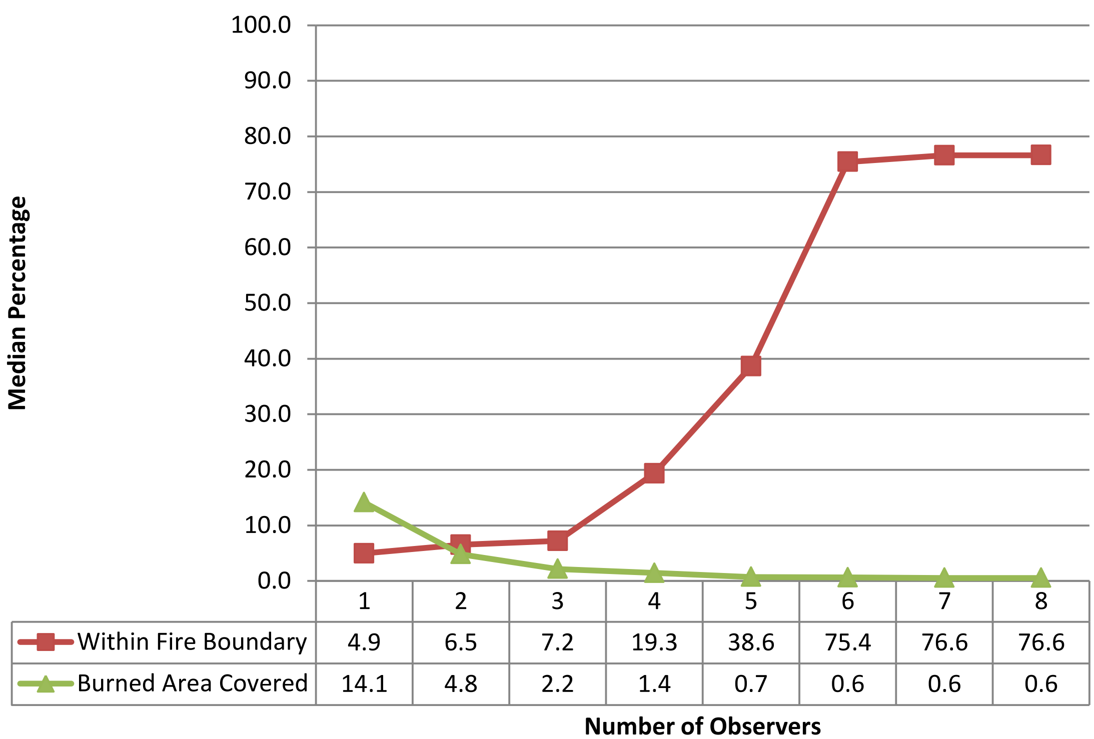

Figure 6 – Effects of Group Size on Fire Visibility

Table 3 – Spatial Distribution and Fire Visibility for each Grouping

Timeframe	Number of Observers	Average Distance (km)	Visible from at Least One Observer (m)	Percentage of Fire Visible from at Least One Observer
2:35 – 2:55pm	2	11.1	1,569.5	15.2%
3:32 – 3:44pm	3	6.2	1,957.7	18.9%
3:44 – 3:56pm	5	7.8	2,179.9	21.1%
4:20 – 4:25pm	2	16.7	562.7	5.4%
4:38 – 4:45pm	2	16.8	1,798.4	17.4%
5:13 – 5:28pm	2	16.6	2,375.6	23.0%
5:28 – 5:36pm	4	7.8	2,763.4	26.7%
6:19 – 6:33pm	4	13.6	2,980.0	28.8%
6:31 – 6:44pm	8	11.0	3,244.9	31.4%
6:53 – 7:07pm	7	6.2	2,316.4	22.4%
7:06 – 7:19pm	4	6.3	2,670.3	25.8%
7:29 – 7:36pm	3	6.9	1,511.3	14.6%
7:44 – 7:54pm	7	9.8	2,527.6	24.4%
8:26 – 8:29pm	2	15.8	1,315.8	12.7%
8:46 – 8:54pm	2	10.5	404.1	3.9%
9:50 – 9:53pm	2	10.3	1,355.4	13.1%
10:11 – 10:16pm	3	6.4	2,277.9	22.0%
10:27 – 10:39pm	5	8.6	2,114.7	20.5%
10:35 – 10:45pm	3	8.2	1,910.5	18.5%
10:57 – 11:10pm	4	7.0	2,469.4	23.9%
11:42 – 11:50pm	2	3.3	2,219.1	21.5%

6. DISCUSSION

6.1 Type of Information in Wildfire Tweets

One of the goals of this research was to determine what kind of information the public disseminates during wildfires. In order to do so, the logical first step was to understand what words were being used in tweets and how often they occurred. After running the dataset through a custom Python code, it was determined that of the 912 tweets sent on June 26th, the most commonly mentioned fire-related keywords were '#WaldoCanyonFire' (141) and 'fire' (84). Considering the high frequency of the official event hashtag (#WaldoCanyonFire), it is safe to say that by the fourth day of the fire, Twitter users were well aware of the hashtag. In fact, of the 83 tweets used in this study, 68 of them used the event hashtag by itself or in combination with other fire-related words. Of course, if the wildfire activity and intensity on June 26th occurred on the first day of the fire, it is likely that an event hashtag would not have been established, making the filtering of tweets more difficult. It is also possible that multiple event hashtags would circulate until an official fire name was determined, which occurred during the Waldo Canyon Fire. On June 23rd, as the public got first wind of the fire, both '#WaldoCanyonFire' and '#PyramidMtnFire' were gaining footing with local Twitter users, with 13 and 26 mentions, respectively. The next day, the fire's official name circulated and '#WaldoCanyonFire' was accepted by the public as the event hashtag, with 74 mentions, compared to 1 mention of '#PyramidMtnFire'.

Other keywords associated with the wildfire event, but not describing the fire activity itself, also garnered frequent mentions. The most frequent of these keywords were variations of ‘evacuation’ (37) and ‘prayers’ (39). While a majority of the evacuation tweets involved users forwarding updates on evacuation areas, a few offered emotional stories of the user or other family members having to evacuate. Tweets offering prayers and condolences, whether they are for victims or first responders, seem to accompany disasters, usually coming from outside the affected area. Of the 39 that came from the affected area, most of them were asking for prayers for not only the Waldo Canyon Fire, but also the entire state, which had seven other ongoing wildfires.

There were a few other noticeable trends in the data as well, the first of which being the tendency for tweets to include photos, especially those related to the event. In fact, of the 912 tweets run through the word count Python code, 319 included hyperlinks to photos. The final dataset of 82 tweets had an even higher percentage, 92.6% or 76 tweets, include photos. Another trend involved the nature of tweets mentioning smoke. Thirty of the 912 tweets mentioned smoke, but none of them made the final dataset due to the fact that they made no mention of fire activity. Instead, all 30 tweets mentioned the irritations accompanied by heavy smoke, including difficulty breathing, the overwhelming smell, and poor visibility.

6.2 Viewsheds and Tracking Wildfires

When the rate of spread of a wildfire increases, so too does the value of fire-related tweets, as they have the potential to fill the time-sensitive void of information that is created by the lagged dissemination of information through official channels. When considering the value of tweets, in general, those that include the geotagged location have significantly greater value than others, as they link detailed coordinates to the content of the tweet. In disasters, including the geotag with a tweet can provide otherwise unknown details to the unfolding story and contribute greatly to situational awareness. Tweets that contribute to situational awareness are defined by Verma et al. (2011) as those that demonstrate an awareness of the scope of the disaster. While the value of a geotagged tweet during a disaster is high, the content can be made even richer by incorporating it into a GIS. When looking at tweets during disasters, especially wildfires, the concern not only lies with what the user is saying, but what they are seeing, as well. Adding tweet locations into a GIS and creating viewsheds allows for the delineation of what is visible or not visible from a user's location, thus making the tweets even more valuable.

For this research, the purpose of adding the viewsheds to the twitter data was to determine if a group of users sending tweets around the same time, and from different locations, could use their unique perspectives to delineate areas that were most likely on fire. Looking back at Figure 5, although the total percentage of the fire covered by the shared viewsheds becomes quite small as more observers are considered, the portions that do fall within the fire boundary account for a much higher percentage of the shared viewsheds. In other words, while less of the fire is visible by all observers as their number increases, the areas that are visible are much more likely to fall within the fire boundary, thus allowing for greater confidence in determining at least a small portion of the fire's location. Naturally, due to high relief and population distribution in the affected area, there were large portions of the fire not visible from any user's location and likely not visible from any point outside of the fire boundary. Looking back at Table 3, many of the observer groups were quite successful in covering a significant portion of the fire with at least one tweet. In fact, considering all 82 tweets throughout the day, just over 32%, or 4,980 acres, of the fire was visible by at least 1 of the observers, most of which was made up of the east-facing canyons and slopes extending down towards the city. Although 10,537 acres of the fire were not visible, the public was likely only concerned with the areas that were visible on the city-facing slopes, as that is when the perceived threat quickly became real.

6.3 User-Generated Information Dissemination

Given the spatial and temporal variation of the tweets in this research, another goal was to determine if users could inform one another of changes in the wildfire activity or location. While it is already known that using Twitter on location-aware devices allows users to function as sensors (Goodchild), it has not been determined if users can also function as fire scouts. Looking closely at the location and timing of each tweet, there were two noticeable ways in which users offered potentially unknown and informative content regarding fire activity. The first considers users that offered unique perspectives of the fire and the second highlights users whose tweets provided powerful content.

In regards to the spatial distribution of the tweets, 75 were sent from a location with longitudinal coordinates further east than the easternmost extent of the fire boundary. While the visibility of the fire on the east-facing slopes varied between these users, they all shared the same perception of the fire activity before it crested the ridge. The remaining seven tweets included the event hashtag and were sent from users offering exclusive views of the fire, either due south of the fire in Manitou Springs or west of the fire in Crystola. All of these tweets were sent during the late morning and early afternoon hours, as conditions were worsening and the fire was beginning to spread more quickly, but had not yet entered Queen's Canyon, located on the west side of ridgeline visible from Colorado Springs. The content within the tweets included descriptions of increasing fire intensity and photos, one in particular showing billowing smoke plumes growing thicker and being blown eastward, indicating the current wind direction and likely direction of spread. While the tweet and included photo were helpful by themselves, the inclusion of the geotag made it possible for anyone to orient the photo and determine where the fire seemed to be spreading. Although there was no guarantee that other users in Colorado Springs saw these tweets, there was at least an effort made by the observers to be informative by offering the first clues of changing fire activity and including the event hashtag.

Users that provide powerful content send tweets that use a combination of clear and concise textual information and a photo that verifies the text. These tweets tend to grab the attention of other users and then circulate to their network of followers via the retweet. In this study, two tweets stood out in the dataset as having a combination of informative textual content and impactful photos that offered clear visibility of the expanding fire boundary. The first tweet, sent at 4:40 pm, included the short message "Crested over the ridge #WaldoCanyonFire" and a photo. While the textual content was concise and somewhat informative, it is the included photo that made this tweet so powerful. In the photo, the statement that the fire had crested the ridgeline is confirmed by a clearly visible eruption of flames on the face of the mountain just above a neighborhood nestled in the foothills. Not 18 minutes later, the second tweet was sent from a location slightly south and much closer to the fire than the first. In addition to an impressive photo, this user managed to cram a great deal of detail into his textual message, stating, “This just became an urban fire. Right behind the MCI building. Wall of Fire. #waldocanyonfire.” Once again, the textual content is verified by the photo, which looks across a small, empty field at the fire encroaching on a building very close to where the user was located. Both of these tweets provided great detail on the fire’s location during a point in the event where the fire was violently spreading, the extent was unknown, and other citizen’s lives and property were suddenly threatened.

7. CONCLUSION

Along the way of developing the methods used in this research, some ways in which it could be altered or expanded upon were identified. For instance, in this research, the effect of flame height on the observer's perspective of the fire was disregarded. This could likely be accounted for by shifting the DEM or elevation of the observer. Secondly, if the temporal resolution of the fire boundary were less coarse, the viewshed calculations could be determined for smaller, shifting boundaries instead of a solitary boundary; however, this alteration would be strictly reliant on more frequent thermal infrared scans during the critical containment periods. Lastly, given that there were a large number viewsheds generated for a short timeframe, the temporal shift in the viewsheds may be worth investigating for any patterns that may emerge.

While the addition of a viewshed certainly makes a tweet more valuable during a disaster, that value is minimal unless the viewshed is utilized during the event. With the quantity of tweets sent during these scenarios, it is not feasible to push tweets through the GIS workflow at an efficient rate without an automated process. Automation of the methods written here could be established through a combination of the Twitter application program interface (API) and Python coding. The application of this process could serve well for wildfires, but given their large scale, it may be better served in other disasters. For example, immediately following an earthquake, tweet viewsheds could be cross-referenced to help identify major damage, fires, or other areas of focus for response. At a localized level, the isolated and immobile nature of an earthquake makes it more likely for the viewhsheds to produce information that is actionable. Theoretically, if the tweets were being monitored by an emergency response agency, and a second perspective of damage was not available, the agency could open a line of communication with that civilian and ask them to provide more information from another view. Of course, there are social implications of the public's awareness of the value of their tweets. If even a few citizens took it upon themselves to tweet from different locations in or around the threat, they could essentially be putting themselves in harm's way. It is worth debating the point at which a tweeting citizen is doing more harm to themselves or the situation than they are benefitting it.

Focusing again on wildfires, with relation to the temporal shifts in viewsheds, is the concept of tracking the shift in tweet locations as the fire grows. For instance, as the fire expands northward, does the location of tweets also shift northward? The temporal tracking of tweet locations as related to a natural hazard has already been done, with one example coming from Mislove et al. (2011), in which they simultaneously animated the progression of the August 23, 2011Virginia earthquake wave and the aggregate count of tweets mentioning "earthquake" per county. The short video shows that in a matter of minutes, some users were tweeting about the earthquake before the shockwave even reached them. Plotting tweet locations in relation to a moving subject may also have applications outside the realm of disasters. Take, for instance, a large sporting event such as the Boston Marathon. This event covers a relatively large area, draws dense crowds of potential tweeters, likely has an event-specific hashtag, and much like a wildfire, has a moving focal point in the runners. By using geotagged tweets along the race route, it seems likely that one could track the progression of the major packs of runners without being in attendance.

The GIS methods described here were developed in an effort to move beyond the accuracy concerns of VGI data and identify a useful application during wildfires. The goal of this research was to enrich the VGI data by adding the viewshed and use it to help track the location of a fast-spreading fire. By separating the tweets into time-based groups and comparing their viewsheds to one another, the research was able to determine possible fire locations with higher confidence than would have been possible without including the viewshed. The research was also able to identify trends in the content of event-related tweets, including the frequency of tweets with photos and the public's quick awareness and acceptance of the event hashtag. The potential for users to inform one another of changes in fire activity or location was also noticed, with tweets offering unique perspectives of the fire or powerful content and photos being the most informative. All in all, it seems more likely that VGI, especially when incorporated into GIS, could help fill the lag in information dissemination during fast-moving events.

REFERENCES

Cameron, P. A., Mitra, B., Fitzgerald, M., Scheinkestel, C. D., Stripp, A., Batey, C., & Cleland, H. (2009). Black saturday: The immediate impact of the February 2009 bushfires in Victoria, Australia. *The Medical Journal of Australia*, *191*(1), 11-16.

Chakraborty, J., Tobin, G.A., & Montz, B.E. (2005). Population evacuation: Assessing spatial variability in geophysical risk and social vulnerability to natural hazards. *Natural Hazards Review*, 1(23), 6.

Corvey, W.J., Vieweg, S., Rood, T., & Palmer, M. (2010). Twitter in mass emergency: What NLP techniques can contribute. *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media*, Los Angeles, California. 23-24.

Cova, T.J. (1999). GIS in emergency management. *Geographical information systems, 2,* 845-858.

Cova, T.J., & Church, R.L. (1997). Modeling community evacuation using GIS. *International Journal of Geographical Information Science, 11,* 763-784.

Cova, T.J., Dennison, P.E., Kim, T.H., & Moritz, M.A. (2005). Setting wildfire evacuation trigger points using fire spread modeling and GIS. *Transactions in GIS, 9,* 603-617.

Crampton, J. W., Graham, M., Poorthuis, A., Shelton, T., Stephens, M., Wilson, M.W., & Zook, M. (2013). Beyond the Geotag: Situating 'big data' and leveraging the potential of the geoweb. *Cartography and Geographic Information Science, 40*(2), 130-139.

Cutter, S.L. (1996). Vulnerability to environmental hazards. *Progress in Human Geography, 20*(4), 529-539.

Cutter, S.L. (2003). GI Science, disasters, and emergency management. *Transactions in GIS, 7*(4), 439-445.

Cutter, S.L., & Emrich, C.T. (2006). Moral hazard, social catastrophe: The changing face of vulnerability along the hurricane coasts. *The ANNALS of the American Academy of Political and Social Science, 604*(1), 102-112.

De Longueville, B., Smith, R. S., & Luraschi, G. (2009). OMG, from here, I can see the flames!: A use case of mining location based social networks to acquire spatio-temporal data on forest fires. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*, 73-80.

Dennison, P.E., Cova, T.J., & Moritz, M.A. (2007). WUIVAC: A wildland urban interface evacuation trigger model applied in strategic wildfire scenarios. *Natural Hazards, 41,* 181-99.

Fisher, P.F. (1991). First experiments in viewshed uncertainty: The accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing, 57,* 1321-1327.

Fisher, P.F. (1992). First experiments in viewshed uncertainty: Simulating the fuzzy viewshed. *Photogrammetric Engineering and Remote Sensing, 58,* 345-352.

Fisher, P.F. (1996). Extending the applicability of viewsheds in landscape planning. *Photogrammetric Engineering and Remote Sensing, 62,* 1297-1302.

Goodchild, M.F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal, 69*(4), 211-221.

Goodchild, M.F., & Glennon, A. (2010). Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth, 3*(3), 231-241.

Goodchild, M.F., & Lee, J. (1989). Coverage problems and visibility regions on topographic surfaces. *Annals of Operations Research, 18,* 175-186.

Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., & Boyd, D. (2011). The revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International Journal of Communication, 5,* 1375-1405.

Maloy, M.A., & Dean, D.J. (2001). An accuracy assessment of various GIS-based viewshed delineation techniques. *Photogrammetric Engineering and Remote Sensing, 67*(11), 1293-1298.

Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter under crisis: Can we trust what we RT?. In the *1st Workshop on Social Media Analytics*, Washington, DC.

Messner, F., & Meyer, V. (2005). Flood damage, vulnerability and risk perception: Challenges for flood damage research. *UFZ-Diskussionspapiere*, No. 13/2005.

Midkiff, S.F., & Bostian, C.W. (2002). Rapidly-deployable broadband wireless networks for disaster and emergency response. In *The First IEEE Workshop on Disaster Recovery Networks (DIREN 2002)*, New York City, NY.

Mills, J. W. (2008). Understanding disaster: GIScience contributions in the ongoing recovery from Katrina. *Transactions in GIS, 12,* 1–4.

Mislove, A., Lehmann, S., Ahn, Y., & Kliman-Silver, C. (2011). *TweetQuake*. Retrieved from: http://blogs.iq.harvard.edu/netgov/2011/08/tweetquake.html

Morrow, B.H. (1999). Identifying and mapping community vulnerability. *Disasters, 23*(1), 1-18.

Ormsby, T., & Alvi, J. (1999). *Extending ArcView GIS*. Redlands: Environmental Systems Research Institute, Inc.

Palen, L., Starbird, K., Vieweg, S., & Hughes, A. (2010). Twitter-based information distribution during the 2009 Red River Valley Flood threat. *Bulletin of the American Society for Information Science and Technology, 36*(5), 13-17.

Pompa-Garcia, M., Solis-Moreno, R., Rodriguea-Tellez, E. Pinedo-Alvarez, A. Avila-Flores, D., Hernandez-Diaz, C., & Velasco-Bautista, E. (2010). Viewshed analysis for improving the effectiveness of watchtowers, in the North of Mexico. *The Open Forest Science Journal, 3,* 17-22.

Pultar, E., Raubal, M., Cova, T.J., & Goodchild, M.F. (2009). Dynamic GIS case studies: Wildfire evacuation and volunteered geographic information. *Transactions in GIS, 13,* 85-104.

Sarcevic, A., Palen, L., White, J., Starbird, K., Bagdouri, M., & Anderson, K. (2012). “Beacons of Hope” in decentralized coordination: Learning from on-the-ground medical twitterers during the 2010 Haiti Earthquake. In the *2012 ACM Conference on Computer Supported Cooperative Work,* Bellevue, WA.

Senaratne, H., Bröring, A., & Schreck, T. (2013). Using reverse viewshed analysis to assess the location correctness of visually generated VGI. *Transactions in GIS, 17,* 369–386.

Showalter, P., & Lu, Y. (2010). *Geospatial techniques in urban hazard and disaster analysis*. New York: Springer.

Simmons, K. M., & Sutter, D. (2012). The 2011 tornadoes and the future of tornado research. *Bulletin of the American Meteorological Society*, *93*(7), 959-961.

Smith, K. (2013). *Environmental hazards: Assessing risk and reducing disaster.* New York: Routledge.

Sorensen, J. H. (2000). Hazard warning systems: Review of 20 years of progress. *Natural Hazards Review*, *1*(2), 119-125.

Starbird, K., Muzny, G., & Palen, L. (2012). Learning from the crowd: Collaborative filtering techniques for identifying on-the-ground twitterers during mass disruptions. In the *Proceedings of the Conference on Information Systems for Crisis Response and Management (ISCRAM 2012)*, Vancouver, BC, Canada.

Starbird, K., & Palen, L. (2011). “Voluntweeters”: Self-organizing by digital volunteers in times of crisis. In the *2011 Conference on Computer Human Interaction (CHI 2011)*, Vancouver, BC, Canada.

Starbird, K., & Palen, L. (2012). (How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising. In the *2012 ACM Conference on Computer Supported Cooperative Work*, Bellevue, WA.

St. Denis, L.A., Hughes, A., & Palen, L. (2012). Trial by fire: The development of trusted digital volunteers in the 2011 Shadow Lake Fire. In the *Proceedings of the Conference on Information Systems for Crisis Response and Management (ISCRAM 2012)*, Vancouver, BC, Canada.

Sutton, J., Palen, L., & Shklovski, I. (2008). Backchannels on the front lines: Emergent uses of social media in the 2007 Southern California wildfires. In the *Proceedings of the 5th International ISCRAM Conference*, Washington, D.C.

Verma, S., Vieweg, S., Corvey, W.J., Palen, L., Martin, J.H., Palmer, M., Schram, A., & Anderson, K.M. (2011). Natural language processing to the rescue?: Extracting "situational awareness" tweets during mass emergency. In the *Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, Spain.

Vieweg, S., Hughes, A.L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: What Twitter may contribute to situational awareness. In the *CHI 2010: Crisis Informatics*, Atlanta, GA.

Wisner, B. (Ed.). (2004). At risk: natural hazards, people's vulnerability and disasters. *Psychology Press*.

Wood, N.J., & Good, J.W. (2004). Vulnerability of port and harbor communities to earthquake and tsunami hazards: The use of GIS in community hazard planning. *Coastal Management, 32,* 243-269.

Zlatanova, S., & Li, J. (2008). *Geospatial Information Technology for Emergency Response*. London: Taylor & Francis Group.

Zook, M., Graham, M., Shelton, T., & S. Gorman. (2010). Volunteered geographic information and crowdsourcing disaster relief: A case study of the Haitian earthquake. *World Health and Medical Policy, 2*(2), 7-33.